

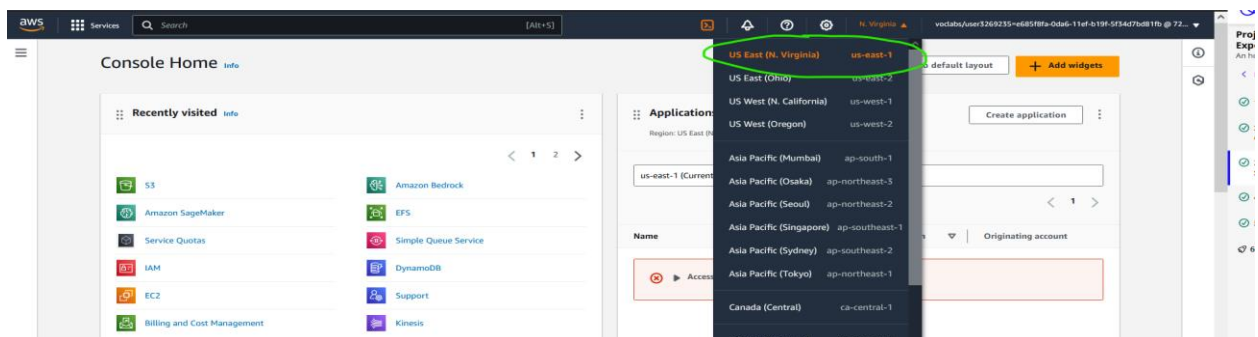
ENVIRONMENT SETUP FOR SAGEMAKER NOTEBOOK INSTANCE

1. Log in to Your AWS Account

- ✓ Open your web browser and go to the [AWS Management Console](https://aws.amazon.com/console/).
- ✓ Enter your credentials to log in.

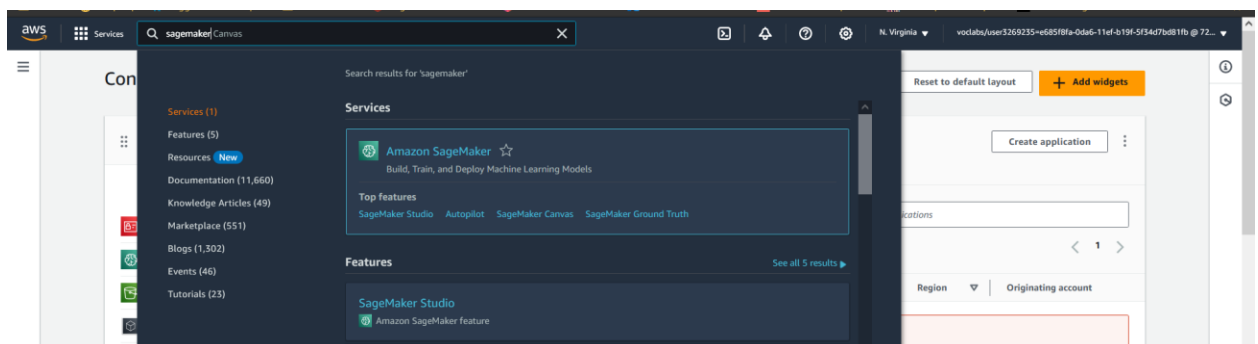
2. Select Your Region

- ✓ Once logged in, locate the region selector in the top right corner of the page.
- ✓ Click on the dropdown menu and choose your desired region. For example, select "US-East" if that is your preferred region.
- ✓ Ensure you consistently work within this region throughout the setup process.



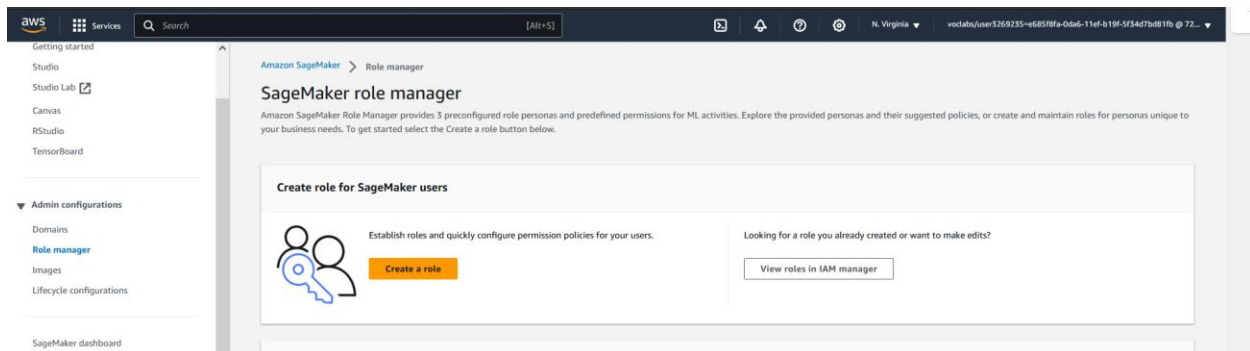
3. Navigate to Amazon SageMaker

- ✓ In the AWS Management Console, find the search bar at the top of the page.
- ✓ Type "SageMaker" into the search bar.
- ✓ From the search results, click on "Amazon SageMaker" to navigate to the SageMaker homepage.



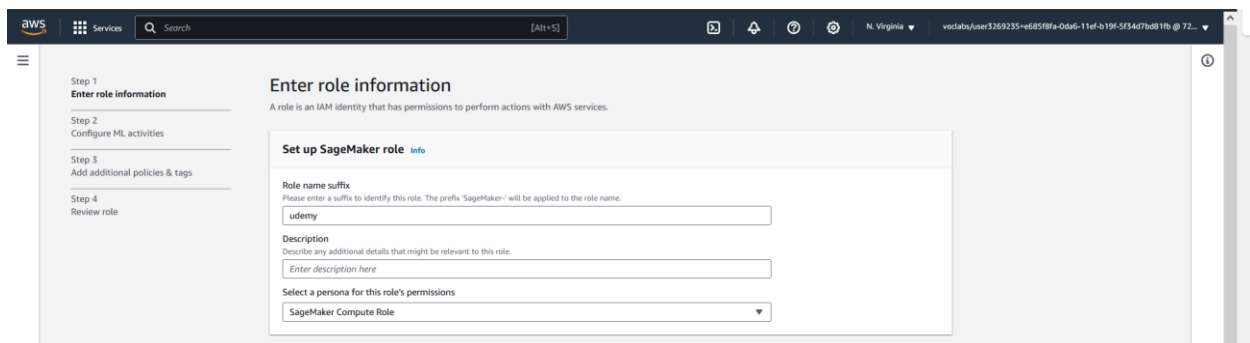
4. Access and Configure Admin Settings

- ✓ On the SageMaker homepage, look for the vertical icon menu on the left side of the page.
- ✓ Click on "Admin Configurations" to expand the dropdown menu.
- ✓ From the dropdown list, select "Role Manager."
- ✓ Click on the "Create a Role" tab to begin creating a new role.



5. Create a New Role

- ✓ In the "Create a Role" page, enter a name of your choice for the "Role Name Suffix" field.
- ✓ In the "Select a persona for this role's permissions" section, choose "SageMaker Compute Role."
- ✓ Click the "Next" button to proceed.



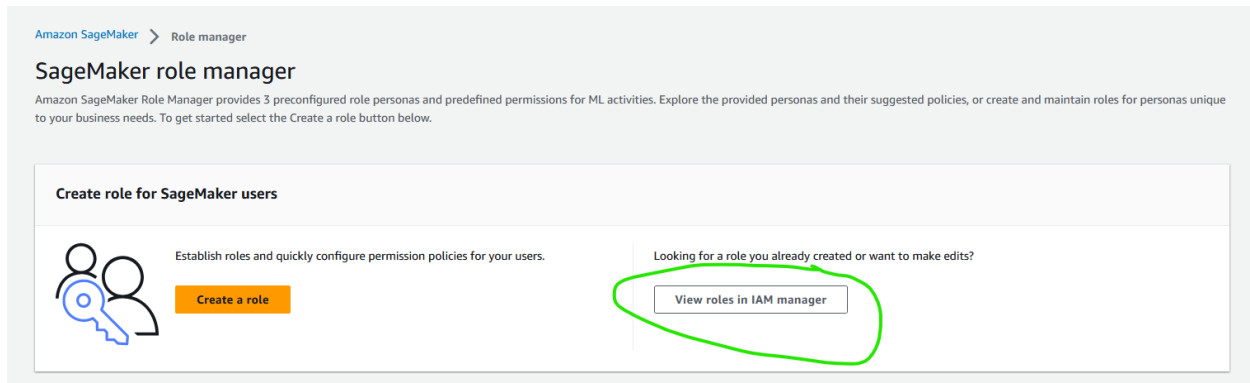
6. Configure Role Permissions

- ✓ On the next page, you will see a list of permissions.
- ✓ Refer to the provided figure to identify the 8 specific items to select.
- ✓ Select only those 8 highlighted items.
- ✓ Click the "Next" button to continue.

<input type="checkbox"/>	Access Required AWS Services	Permissions to access S3, ECR, Cloudwatch and EC2. Required for execution roles for jobs and endpoints.
<input checked="" type="checkbox"/>	Run Studio Applications	Permissions to operate within a Studio environment. Required for domain and user-profile execution roles.
<input checked="" type="checkbox"/>	Manage ML Jobs	Permissions to manage SageMaker jobs across their lifecycles.
<input checked="" type="checkbox"/>	Manage Models	Permissions to manage SageMaker models and Model Registry.
<input checked="" type="checkbox"/>	Manage Pipelines	Permissions to manage SageMaker Pipelines and pipeline executions.
<input checked="" type="checkbox"/>	Search and visualize experiments	Permissions to audit, query lineage and visualize SageMaker Experiments.
<input checked="" type="checkbox"/>	Manage Model Monitoring	Permissions to manage monitoring schedules for SageMaker Model Monitor.
<input checked="" type="checkbox"/>	S3 Full Access	Permissions to perform all S3 operations
<input type="checkbox"/>	S3 Bucket Access	Permissions to perform operations on specified buckets.
<input type="checkbox"/>	Query Athena Workgroups	Permissions to execute and manage Amazon Athena queries.
<input type="checkbox"/>	Manage Glue Tables	Permissions to create and manage Glue tables for SageMaker Feature Store and Data Wrangler.
<input type="checkbox"/>	Canvas Core Access	Permissions to perform experimentation in Canvas (i.e, basic data prep, model build, validation)
<input type="checkbox"/>	Canvas Data Preparation (powered by Data Wrangler)	Permissions to perform end-to-end data preparation in Canvas (i.e, aggregate, transform and analyze data, create and schedule data preparation jobs on large datasets).
<input type="checkbox"/>	Canvas AI Services	Permissions to access ready-to-use models from Amazon Bedrock, Amazon Textract, Amazon Rekognition, and Amazon Comprehend. Additionally, user can fine-tune foundation models from Amazon Bedrock and Amazon SageMaker Jumpstart.
<input type="checkbox"/>	Canvas MLOps	Permission for Canvas users to directly deploy model to endpoint.
<input type="checkbox"/>	Canvas Kendra Access	Permission for Canvas to access Kendra for enterprise document search. The permission is only given to your selected index names in Kendra.
<input checked="" type="checkbox"/>	Use MLflow	Permissions to manage experiments, runs, and models in MLflow.
<input type="checkbox"/>	Manage MLflow Tracking Servers	Permissions to manage, start, and stop MLflow Tracking Servers.
<input type="checkbox"/>	Access required to AWS Services for MLflow	Permissions for MLflow Tracking Servers to access S3, Secrets Manager, and Model Registry.

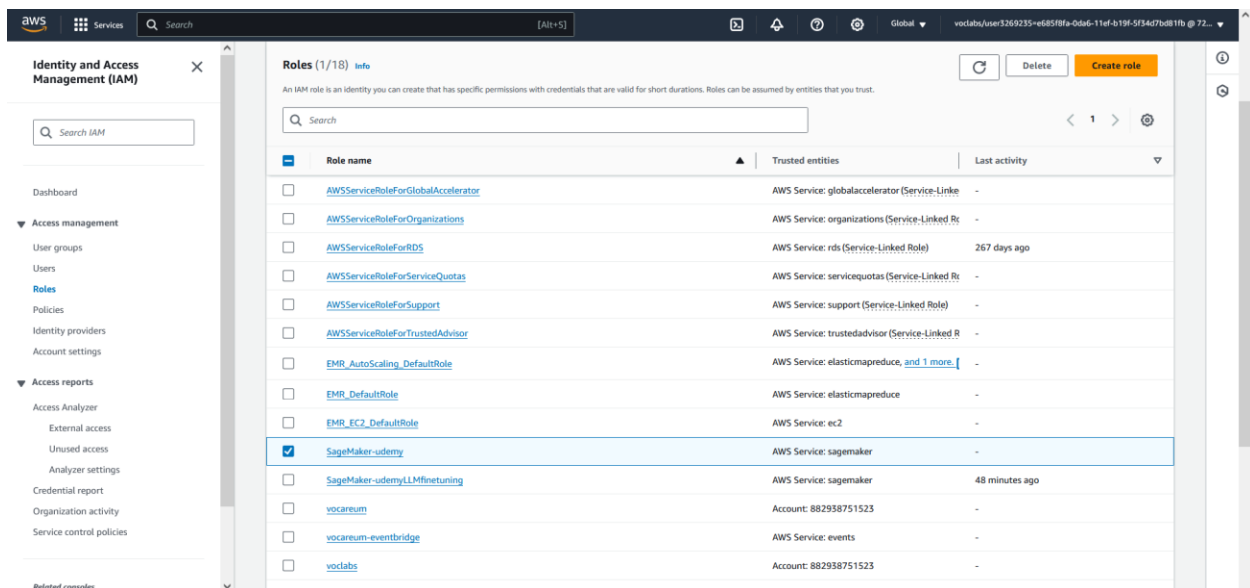
7. Finalize Role Creation

- ✓ On the subsequent page, review the information and click "Next."
- ✓ Click "Submit" to complete the role creation process.
- ✓ After successful submission, you will be redirected back to the SageMaker Role Manager page.
- ✓ On this page, click on "View roles in IAM manager" to open a new tab displaying the list of available roles.



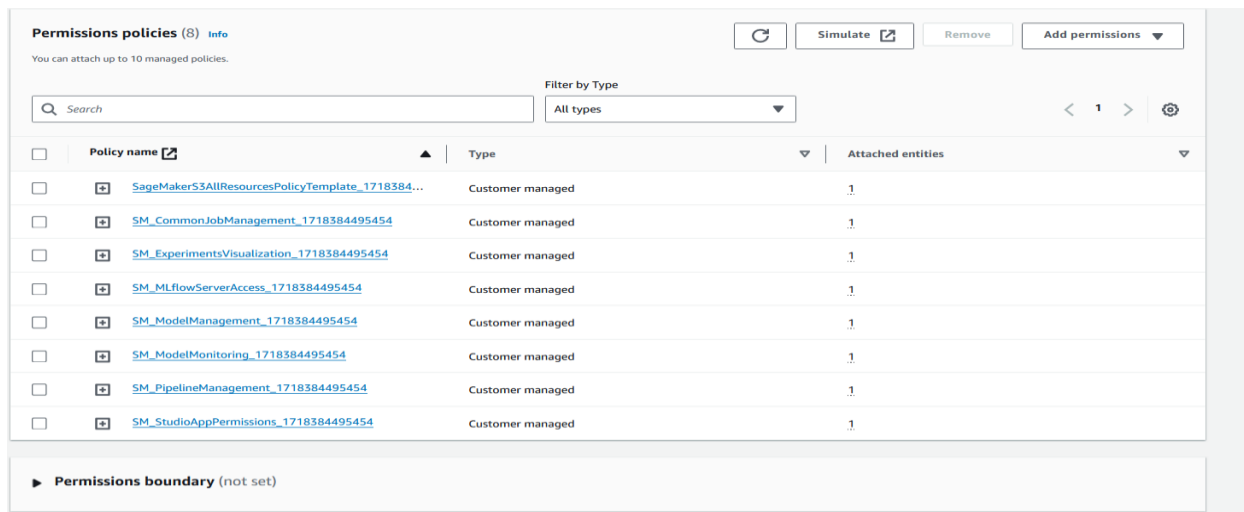
8. Locate the Newly Created Role

- ✓ In the IAM manager tab, you will see a list of available roles.
- ✓ Use the search bar to find the role you just created. Typically, it will have a prefix starting with "SageMaker" followed by the name you assigned.
- ✓ Enter the prefix and the name to locate your role quickly.



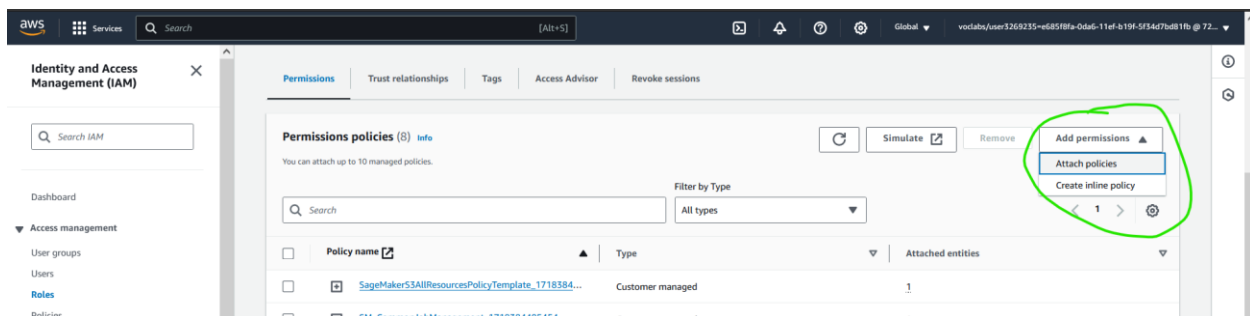
9. View Role Policies

- ✓ From the list of roles in the IAM manager, click on the role you just created.
- ✓ This will open a detailed view of the role, displaying all the policies that have been added to it.



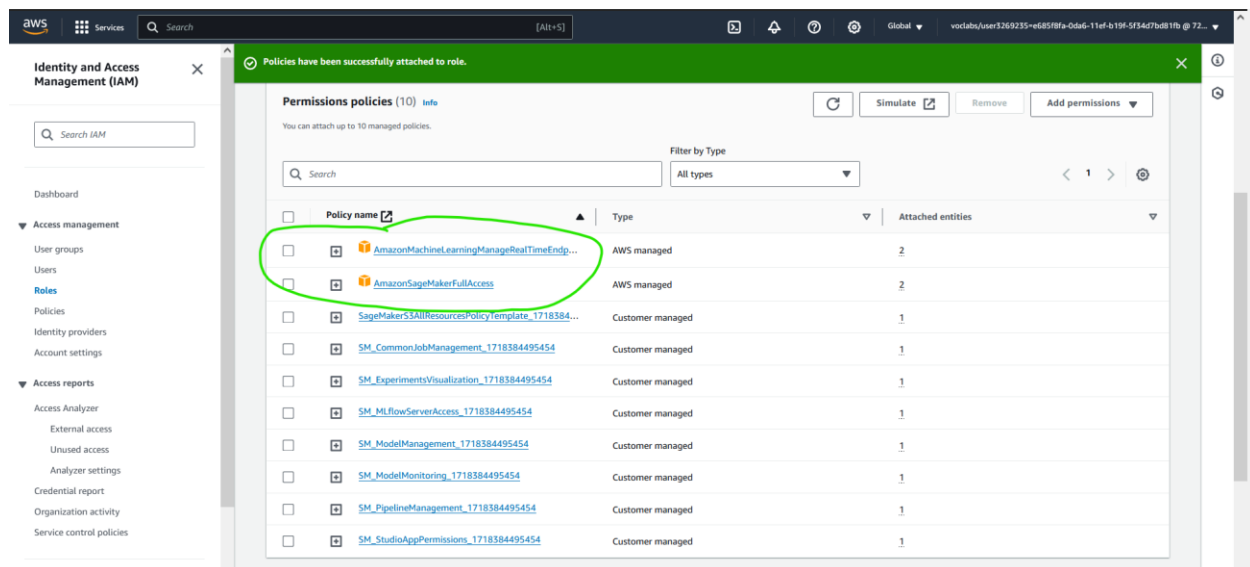
10. Add Additional Permissions

- ✓ In the detailed view of the role, locate the "Add permissions" button in the right-hand corner.
- ✓ Click on "Add permissions" and select "Attach policies" from the dropdown menu.



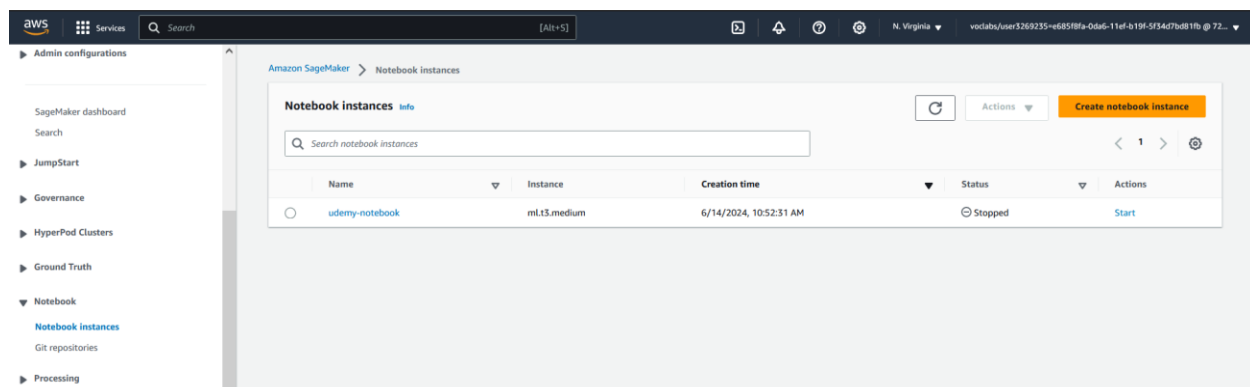
11. Attach Additional Policies

- ✓ In the "Attach policies" screen, navigate to the "Other permission policies" section.
- ✓ Use the search bar to find and select the two highlighted permissions as indicated.
- ✓ After selecting the policies, scroll down and click on "Add permissions" located below the list to attach the permissions to the role.



12. Create a Notebook Instance

- ✓ Navigate back to the SageMaker home page.
- ✓ On the left-hand side menu, locate and click on "Notebook."
- ✓ From the dropdown options, click on "Notebook instances."
- ✓ On the right-hand side of the page, a new section will appear.
- ✓ In the top right-hand corner of this section, click on "Create notebook instance."



13. Fill Out Notebook Instance Details

- ✓ Fill in the required information as follows:
- ✓ Notebook instance name: Enter a name for your notebook instance.
- ✓ Notebook instance type: Select the appropriate instance type based on your requirements.
- ✓ IAM role: Choose the role you created earlier from the dropdown list under "Permissions and encryption."
- ✓ Encryption key: Leave this as default unless you have specific encryption requirements.

Ensure all details are correctly filled before proceeding with creating the notebook instance.

Notebook instance name

my-udemy-notebook

Maximum of 63 alphanumeric characters. Can include hyphens (-), but not spaces. Must be unique within your account in an AWS Region.

Notebook instance type

ml.t3.medium

Platform identifier [Learn more](#)

Amazon Linux 2, Jupyter Lab 3

► Additional configuration

Permissions and encryption

IAM role

Notebook instances require permissions to call other services including SageMaker and S3. Choose a role or let us create a role with the [AmazonSageMakerFullAccess](#) IAM policy attached.

SageMaker-udemy

Create role using the role creation wizard

Root access - optional

☒ Enable - Give users root access to the notebook

☐ Disable - Don't give users root access to the notebook

Lifecycle configurations always have root access

Encryption key - optional

Encrypt your notebook data. Choose an existing KMS key or enter a key's ARN.

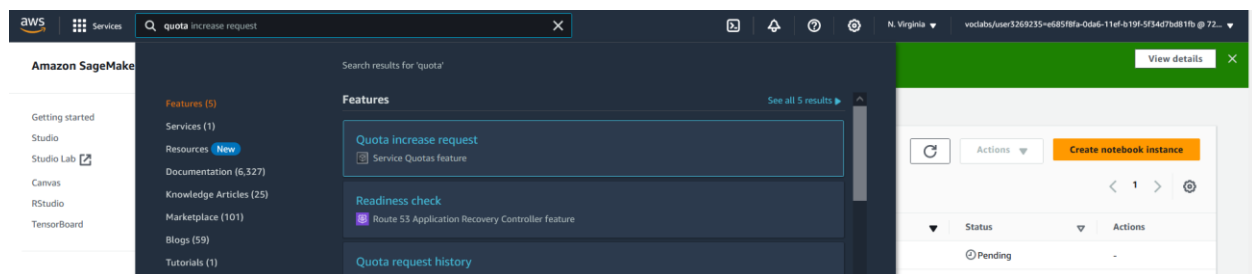
No Custom Encryption

14. Create Notebook Instance

- ✓ - After filling out the required information, click on "Create notebook instance" located at the bottom right corner of the page.
- ✓ - Wait for the notebook instance to be created.

15. Request GPU Instance for Training

- ✓ - Navigate to the AWS Management Console.
- ✓ - In the search bar at the top, type "quota" and select "Quota increase request" from the dropdown list that appears.



16. Request GPU Instance Quota Increase

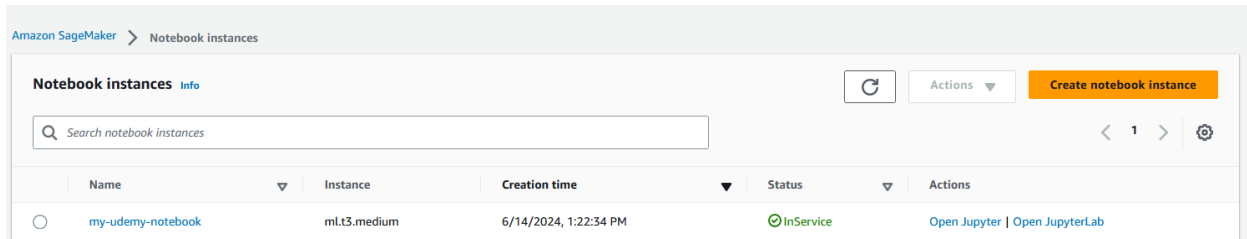
- ✓ - In the AWS Management Console, navigate to the AWS Services dropdown menu at the top of the page or use the search bar to find "Amazon SageMaker".
- ✓ - Click on "View quotas" in the SageMaker service page.
- ✓ - In the "Search by quota name" box, enter "ml.g5.2xlarge".
- ✓ - Locate the row corresponding to "ml.g5.2xlarge" for training job usage and select the circle next to it.
- ✓ - Click on the "Request increase at account-level" button at the top of the page.
- ✓ - In the request form, enter "1" instance.
- ✓ - Submit the request.

17. Monitor Notebook Instance Creation

- ✓ - Return to the SageMaker console and monitor the status of the notebook instance you created.
- ✓ - Once the status changes to "InService", the notebook instance is ready for use.

18. Start LLM Fine-tuning

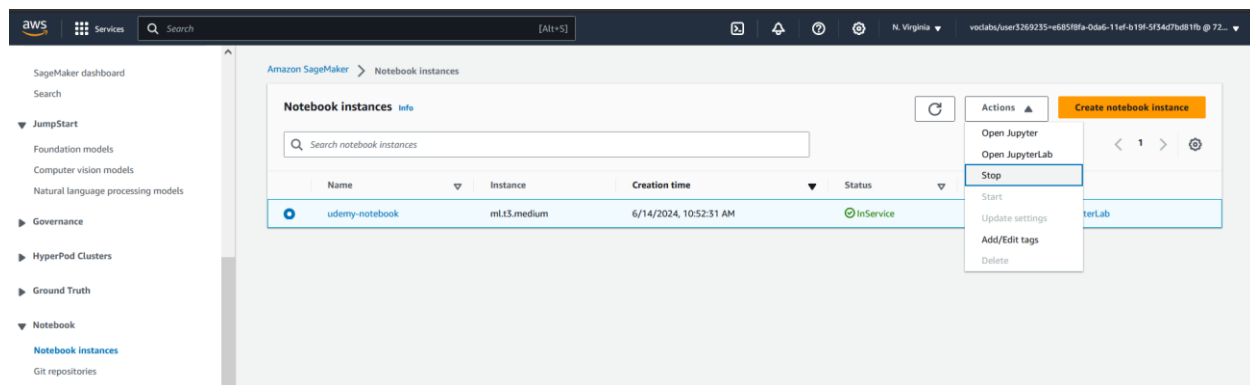
- ✓ Open the Jupyter notebook associated with your SageMaker instance.
- ✓ Start your Language Model (LLM) fine-tuning process as planned.



19. Manage Notebook Instance

- ✓ To perform actions on your notebook instance, locate the instance in the SageMaker console.
- ✓ Use the Action icon (typically represented by three vertical dots or a gear icon) next to the notebook instance to:
- ✓ Start: Click to start the notebook instance.
- ✓ Stop: Click to stop the notebook instance when not in use.
- ✓ Update: Click to modify the instance settings or configurations.
- ✓ Delete: Click to remove the notebook instance when it's no longer needed.

These actions allow you to efficiently manage the lifecycle of your SageMaker notebook instance based on your requirements.



Summary:

1. Log in to AWS Management Console.
2. Select region (e.g., US-East) from top right corner.
3. Search "SageMaker" in AWS console and click to open SageMaker.
4. Click "Admin Configurations" > "Role Manager" > "Create a role" tab.
5. Enter Role Name Suffix, select "SageMaker Compute Role," click "Next."
6. Select 8 highlighted permissions, click "Next," then "Submit."
7. Click "View roles in IAM manager" to see created roles.
8. Search for role starting with "SageMaker" prefix.
9. Click role to view attached policies.
10. Click "Add permissions" > "Attach policies."
11. Search and select highlighted permissions, click "Add permissions."
12. Go to SageMaker home, click "Notebook" > "Notebook instances."
13. Click "Create notebook instance," fill details including new role.
14. Click "Create notebook instance" to start creation.
15. Search "quota" > "Quota increase request" > select "View quotas."
16. Search "ml.g5.2xlarge," select circle, request 1 instance, submit.
17. Monitor notebook status until "InService."
18. Open Jupyter notebook to begin LLM fine-tuning.
19. Use Action icon to start, stop, update, or delete the notebook instance.